Compare two problems below: what parameter we intend to estimate?

I. Sample of 400 randomly selected flight reservations was chosen. 10 of them were canceled on the day of the flight. Find a 98% confidence interval for the true percentage of flight reservations being canceled on the day of the flight.

Hint: ***Qualitative Data*** **Parameter:** **P**

II. Sample of 50 days was randomly chosen and percentage of canceled flight reservations was recorded for each day. Find a 94% confidence interval for the average percentage of flight reservations being canceled on the day of the flight.

Hint: ***Quantitative Data*** **Parameter:** **μ**

Confidence Interval for a Population Mean

1. **A physician wanted to estimate the mean length of time μ that a patient had to wait to see him after arriving at the office. A random sample of 36 patients showed a mean waiting time of 23.4 minutes and a standard deviation of 7.2 minutes. Find a 96% confidence interval for μ.**

Parameter of interest is Population Mean (μ)

Point – estimator is sample mean**:** $\bar{x} = 23.4 \text{ min}$

To find critical value $Z_{\alpha/2}$ take half of confidence level (.96:2 = .48), then use Normal table to locate entry closest to 0.4800----> 0.4798, which corresponds to Z-score = 2.05

CI: $\bar{X} \pm Z_{\alpha/2} \frac{S}{\sqrt{n}} = 23.4 \pm 2.05 \frac{7.2}{\sqrt{36}}$ **= 23.40 ± 2.46** **SE = 2.46**

Conclusion: We are 96% confident that average time to wait in physician office is within (22.94, 25.86) min

2. The quality control officer collects a random sample of 49 yardsticks from the day's production run. The sample mean is 36.00 inches and the standard deviation is 1.4 inches. Find a 99% confidence interval for the mean length of all yardsticks made that day.

Parameter of interest is Population Mean (μ)

Point – estimator is sample mean**:** $\bar{x} = 36 \text{ in}$

To find critical value $Z_{\alpha/2}$ take half of confidence level (.99:2 = .4950), then use Normal table to locate entry closest to 0.4950---> 0.4945 and 0.4955, compute average between Z = 2.57 and Z = 2.58 which corresponds to Z-score = 2.575

CI: $\bar{X} \pm Z_{\alpha/2} \frac{S}{\sqrt{n}} = 36.0 \pm 2.575 \frac{1.4}{\sqrt{49}}$ **= 36 ± .52** **SE = .52**

Conclusion: We are 99% confident that average length of all yardsticks made that day is within (35.48, 36.52) in.

3. During the last Super-Bowl Sunday, Adrian and his buddies ordered 27 pizzas from *Pizzas To Go* at different times. The average delivery time proved to be 23.7 min, with a standard deviation of 10.7 min. Assume a normal distribution. Feeling this was far too long a delay, Adrian and his friends decided to buy the *28th* pizza elsewhere if it appeared that the delivery time for *Pizzas To Go* exceeded 30 min. Set alpha at 1%. Will they order elsewhere?

Hypothesized Parameter value (Population Mean): μ_o = 30 min

Point –estimator (sample mean): $\bar{x} = 23.7$ min

CI: $\bar{X} \pm t_{\alpha/2} \frac{S}{\sqrt{n}} = 23.7 \pm 2.779 \frac{10.7}{\sqrt{27}} = \mathbf{23.7 \pm 5.72}$

Answer: (17.98, 29.42) min. Interpretation: You can be 99% confident that the pizzas from Pizzas To Go will arrive in less than 30 minutes, since 30 does not fall in the interval. Continue ordering from Pizzas To Go.

4. Sample of twenty five 16-ounce jars of Hot Sauce averages 15.7 ounces with standard deviation of 0.96 ounces. At the 95 percent level of confidence, can you assert the jars are filled with a mean of 16 ounces? Would you suggest that company needs to put more hot sauce into the jars to insure labeling accuracy?

Hypothesized Parameter value (Population Mean): μ_o = 16 oz

Point –estimator (sample mean): $\bar{x} = 15.7$ oz

CI: $\bar{X} \pm t_{\alpha/2} \frac{S}{\sqrt{n}} = 15.7 \pm 2.064 \frac{0.96}{\sqrt{25}} = \mathbf{15.7 \pm 0.40}$ **SE = 0.40 oz, μ: (15.3, 16.1) oz**

Conclusion: Based on constructed CI the average content of Hot Sauce jars is within (15.3, 16.1) oz, which means we don't have sufficient evidence to prove that average content of Hot Sauce jar is equal to 16 oz.

5. A 94 % confidence interval for the mean amount of coffee dispensed by vending machine is: (7.2, 7.6) oz.

a) Interpret this confidence interval in context of the problem.

We are 94% confident that average amount of coffee dispensed by vending machine is within (7.2, 7.6) oz.

b) What is the value of the **point estimator** that was used to construct this interval? $\bar{x} = 7.4$ oz

c) Calculate the **sampling error** to within the average length of time was estimated. SE = 0.2 oz

d) Suppose you want **to reduce the width of the confidence interval to half of its current size**. Calculate the approximate sample size to achieve the desired accuracy of the estimate. It is known that observations ranged in value between **6.8 oz** and **8.8 oz**. Assume that the **range** of the data is equal to **4 σ**.

Range = 8.8 – 6.8 = 2.0 oz, R = 4 σ = 2.0 oz □ σ = 0.5 oz; new reduced SE = .2 /2 = 0.1 oz

$$n = \left(\frac{Z_{\alpha/2} \cdot \sigma}{SE} \right)^2 = \left(\frac{1.88 \cdot 0.5}{0.1} \right)^2 = 88.36$$

Answer: n = 89 (Round up!)

6. Fat Harry's, a popular student hangout, sells 16-ounce glasses of beer. Ten students purchase a total of 22 glasses and, using their own measuring cup, estimate the mean contents. The sample mean is 15.2 oz with S = 0.86 oz. At the 95% level of confidence, are the students getting their money's worth?

Hypothesized Parameter value: $\mu_o = 16$ oz　　　　**Point – estimator:** $\bar{x} = 15.2$ oz

CI: $\bar{X} \pm t_{\alpha/2} \frac{S}{\sqrt{n}}$ **= 15.2 ±** $2.08 \frac{0.86}{\sqrt{22}}$ **= 15.2 ± 0.4**　　**SE = 0.40 oz**

Answer: (14.8, 15.6) oz. You are 95% confident that these students are not getting the 16 oz of beer (too much foam), since the 16 oz does not fall within the interval. There needs to be a boycott of this pub or *at* least an organized picket. Demand your rights to a full mug!

7. Suppose a 90% confidence interval for μ turns out to be (110, 260). Based on the interval, do you believe the average is equal to 270?

A) Yes, and I am 100% sure of it. B) Yes, and I am 90% sure of it. **C) No, and I am 90% sure of it**. D) No, and I am 100% sure of it.

8. Suppose a 98 % confidence interval for μ turns out to be (1000, 2100). If this interval was based on a sample of size n = 19 explain what assumptions are necessary for this interval to be valid?

A) The sampling distribution must be biased with 18 degrees of freedom.

B) The sampling distribution of the sample mean must have a normal distribution.

C) The population must have an approximately normal distribution.

D) The population of salaries must have an approximate t distribution.

9. A 95% confidence interval for a population mean is (28, 35). Can you reject the null hypothesis that μ=36 at 5 % significance level? Why?

Yes, Ho: μ=36 could be rejected at 5 % significance level since CI doesn't include μ=36 as a plausible parameter value.

10. Suppose a 90% confidence interval for μ turns out to be (14, 66). To make more useful inferences from the data, it is desired to reduce the width of the confidence interval. How it can be possible?

a. Increasing the confidence level　　c. a and b
b. Decreasing the sample size　　**d. Neither a nor b.**

11. Two students are doing a statistics project in which they drop toy parachuting soldiers off a building and try to get them to land in a hula-hoop target. They count the number of soldiers that succeed and the number of drops total. In a report analyzing their data, they write the following:
"We constructed a 95% confidence interval estimate of the proportion of jumps in which the soldier landed in the target, and we got [0.50, 0.81]. We can be 95% confident that the soldiers landed in the target between 50% and 81% of the time. Because the army desires an estimate with greater precision than this (a narrower confidence interval) we would like to repeat the study with a larger sample size, or repeat our calculations with a higher confidence level."
How many errors can you spot in the above paragraph?

Answer: There are three incorrect statements:

First, the first statement should read "…the proportion of jumps in which soldiers land in the target." (We're estimating a population proportion.)
Second, the second sentence also refers to past tense and hence implies sample proportion rather than population proportion. It should read, "We can be 95% confident that soldiers land in the target between 50% and 81% of the time." (The difference is subtle but shows a student misunderstanding.)
And the third error is in the last sentence. A higher confidence level would produce a wider interval, not a narrower one.

12. A company needs to estimate the average total compensation of CEOs in the service industry. Data were randomly collected from 35 CEOs and the 95% CI was calculated to be ($2,256,000, $5,580,000). What assumptions are necessary for this CI to be valid?

1. The sample is randomly selected from a population of total compensations that is a t distribution.
2. **None. The Central Limit Theorem applies.**
3. The total compensation of CEOs in the service industry is approximately normally distributed.
4. The distribution of the means is approximately normal.

13. We have calculated a confidence interval based on a sample of size n = 100. Now we want to get a better estimate with a margin of error that is only one-third as large. How large does our new sample need to be?

A) 200 **B) 900** C) 25 D) 400

Confidence Interval for a Population Proportion

1. The Washington Post reported that 90% of all high-school students had computers at their home. If a sample of 1,020 students is taken, which reveals that 873 students have computers at home; does a 99% confidence interval support The Washington Post?

Hypothesized Parameter value (Population Proportion): $p_o = 90\%$

$\hat{p} = \frac{x}{n} = \frac{873}{1020} = 0.86$ **CI:** $\hat{p} \pm Z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}} = 0.86 \pm 2.575\sqrt{\frac{0.86 \cdot 0.14}{1020}} = 0.86 \pm 0.03$ **Answer: CI (0.83, 0.89)**

Confidence interval doesn't support The Washington Post since $p_o = 90\%$ is outside the constructed CI.

2. A union official wanted to get an idea of whether a majority of workers at a large corporation would favor a contract proposal. She surveyed 500 workers and found that 240 did not favor the proposal.

(a) Find a 96% confidence interval for the proportion of all the workers who favor the contract proposal.

$\hat{p} = \frac{x}{n} = \frac{260}{500} = 0.52$ **CI:** $\hat{p} \pm Z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}} = 0.52 \pm 2.05\sqrt{\frac{0.52 \cdot 0.48}{500}} = 0.52 \pm 0.05$

(b) Find the maximum error (sampling error) of the estimate.

SE = 0.05

(c) Based on the results of part (a), can we conclude that the contract will be ratified by the membership?

CI: (0.47, 0.57) We cannot be sure since the proportion of all the workers who favor the proposal could be below than 50%.

3. To estimate the proportion p of passengers who had purchased tickets for more than $400 over a year's time, an airline official obtained a random sample of 75. The number of those purchasing tickets for more than $400 was 45.

(a) What is a point estimate for *p*?

$$\hat{p} = \frac{x}{n} = \frac{45}{75} = 0.60$$

(b) Find a 90% confidence interval for *p*.

CI: $\hat{p} \pm Z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}} = 0.60 \pm 1.645\sqrt{\frac{0.60 \cdot 0.40}{75}} = 0.60 \pm 0.09$ **SE = 0.09**

Interpretation: (0.51, 0.69) We are 90% confident that true proportion of passengers who had purchased tickets for more than $400 over a year's time is between 51% and 69%.

(c) How many passengers are required in the sample to reduce in half the sampling error within which we need to estimate the proportion of passengers who had purchased tickets for more than $400 over a year's time?

$$n = \frac{\left(z_{\alpha/2}\right)^2 \hat{p}\hat{q}}{(SE)^2} = \frac{(1.645)^2 .60 \cdot .40}{(0.045)^2} = 320.7$$

n = 321

d) After reducing SE in half, provide a new (reduced) CI and interpret it in context of the problem.

Initial CI was: 0.60 ± 0.09, the new CI is: 0.60 ± 0.045 ===➔ (0.555, 0.645)

4. A 95% CI for the true percentage of students getting their PhDs is (1.0%, 5.0%). Random sample of 700 graduates was taken for estimation. Was the sample size a sufficient?

Sufficient. $\hat{p} = 3\%$ **np = 700 x 0.03 = 21 > 15**

nq = 700 x 0.97= 679 > 15

5. A 96% confidence interval for the proportion of registered voters across the U.S.A. who would prefer a " Super morally straight person" to be the president of United States is: (56%, 60%).

a) Find a point estimate for the proportion of voters that prefer a president to be a morally straight person.

$\hat{p}$ **= 58%**

b) What is the sampling error of estimation?

SE = 2%

c) How many registered voters are required in the sample to reduce in half the bound within which we need to estimate the true proportion of voters who prefer a "Super morally straight person" to be the president of United States?

$$n = \frac{\left(z_{\alpha/2}\right)^2 \hat{p}\hat{q}}{(SE)^2} = \frac{(2.05)^2 .58 \cdot .42}{(0.01)^2} = 10{,}237.29$$ **Answer: n = 10,238**

6. A small private college is interested to determine if less than 20% of their current students live off campus and drive to class. The college decided to take a random sample of 30 of their current students to use in the analysis. Is the sample size of n = 30 large enough to use this inferential procedure?

A) No. (Hint: The parameter is population proportion.)
B) Yes, since the central limit theorem works whenever proportions are used.
C) Yes, since $n \geq 30$
D) Yes, since both np and nq are greater than or equal to 15

7. Which of the following statements is true?

I. The center of a confidence interval is a population parameter.
II. The bigger the margin of error, the smaller the confidence interval.
III. The confidence interval is a type of point estimate.
IV. A population mean is an example of a point estimate.

(A) I only (B) II only (C) III only (D) IV only (E) None of the above.